



Active Learning Strategies and Active Control of Complexity Growth in Naming Games

William Schueller, Pierre-Yves Oudeyer

► To cite this version:

William Schueller, Pierre-Yves Oudeyer. Active Learning Strategies and Active Control of Complexity Growth in Naming Games. the 5th International Conference on Development and Learning and on Epigenetic Robotics, Aug 2015, Providence, RI, United States. hal-01202654

HAL Id: hal-01202654

<https://inria.hal.science/hal-01202654>

Submitted on 23 Sep 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Active Learning Strategies and Active Control of Complexity Growth in Naming Games

William Schueller

ENS de Lyon, Département de Physique
INRIA & ENSTA ParisTech
Email: william.schueller@inria.fr

Pierre-Yves Oudeyer

INRIA & ENSTA ParisTech
Email: pierre-yves.oudeyer@inria.fr

Abstract—Naming Games are models of the dynamic formation of lexical conventions in populations of agents. In this work we introduce new Naming Game strategies, using developmental and active learning mechanisms to control the growth of complexity. An information theoretical measure to compare those strategies is introduced, and used to study their impact on the dynamics of the Naming Game.

I. INTRODUCTION

What are the mechanisms that allow language formation, transmission and evolution between individuals? How can linguistic conventions be formed, and what is the dynamics of their evolution, in a population of speakers, without a centralized control? How can these mechanisms be computationally and statistically modeled? If we consider language as a dynamical system associating forms, referents and semantic categories, its lexicon can be represented as a relation map between spaces that are *a priori* not related. For example a part of the spoken words-space maps the meaning-space of the objects surrounding us. But if we take two particular objects (say an apple and a coin), there is *a priori* no reason for those words (“apple” and “coin”) to be linked to those two objects in this order. Language may have developed with those two associations interchanged - that means the word “apple” would refer to the money and “coin” would refer to the fruit. The particular pattern observed is a convention stipulating which words refer to which meanings. If we consider a language at a really early formation stage in a non-centralized population of interacting individuals, any permutation of those associations may be electable as a complete operational language at the end.

Whereas existing work has shown how groups of individuals (e.g. robots or humans) could build shared and self-organized linguistic conventions (associations between words and semantic categories) in simple spaces of word forms and meanings [1], [2], or how a human can teach a few new words corresponding to simple concepts to a robot [3], we are still very far from being able to construct systems that can learn and share efficiently a large number of words corresponding to various kinds of concepts (objects, properties, actions, time, etc.). This is partly due to the fact that when confronted to conceptual or lexical spaces that are large and/or high-dimensional, the mechanisms elaborated so far to infer the meaning of words and to converge towards a shared convention are under-constrained.

A possible solution to this problem is the **developmental approach** [4]–[8]. In particular, a number of arguments in statistical machine learning [9] and developmental sciences [10] indicate that the active control of the complexity of learning situations (based for example on empirical measures of information gain) might permit to dramatically increase the efficiency of the language learning and formation processes. The impact of such developmental mechanisms on learning efficiency, such as active learning mechanisms that regulate the growth of complexity in learning situations, has been so far mostly studied in the context of single robots acquiring sensorimotor skills (e.g. [4], [5], [11]). In this article, we explore how such active learning mechanisms can also be used for and how they can impact the dynamics of models describing the formation of linguistic conventions in a population of agents.

Computational and statistical modeling of semiotic and social dynamics is well-defined in the Naming Games introduced by L. Steels [12]–[14], initially used to describe language formation in a population of agents using spoken words to refer to colored geometric shapes (*The Talking Heads Experiment* [12], [15]). Theoretical work on this framework has also been done, including statistical studies on the Naming Games [14], [16], on the introduction of new available elements in the vocabulary [17], as well as information theoretical analysis of vocabulary patterns [18].

Naming Games are simple interaction models between agents, which each have their own vocabulary, in most cases represented as an association matrix between *words* and *meanings*. One interaction usually involves two randomly picked agents, a *speaker* and a *hearer*. The *speaker* chooses a *word-meaning* association in its vocabulary (possibly inventing a new one), and utters the associated *word*. The *meaning* interpreted by the *hearer* is then compared to the one initially chosen by the *speaker*, and they both eventually update their vocabulary according to the results, that is success or failure, translated by whether the two *meanings* match or not.

The agents involved in the interaction may add the uttered association to their vocabulary or not, remove any synonyms or not, same for homonyms, etc. While in most Naming Game models, speakers choose a meaning randomly, there are other possible alternatives: the speaker can for example choose a known meaning or an unknown one (i.e. associated

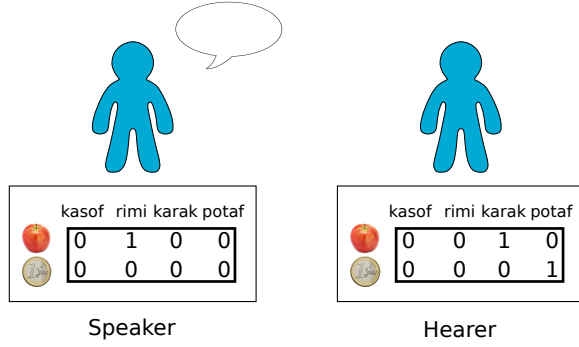


Fig. 1. Example of an interaction in the Naming Game. There are here 2 meanings (apple and coin) and 4 possible words. The speaker can either refer to the apple, by saying the word *rimi*, or invent a new association for the coin. The first case would lead to failure. The second one would lead to success only if the chosen word is *potaf*.

or not to at least one word), or make its choice based on trust levels in the associations, etc. To sum up, a **wide variety of strategies** exist, and some of them may lead to interesting behaviors, exhibiting particular patterns of vocabulary size, distribution and convergence speed at the population level [19]. But those patterns are not easy to represent, or quantify, and a great set of different measures are used in existing work [16], [19], [20], for example the success rate of the interactions, or mean number of associations per agent (proportional to total number of associations in the population), or number of different words used over the whole population, among others.

In this article, we will formulate and experimentally evaluate several active strategies that allow the speaker to decide which meaning to use in each round of the naming game, and in particular to decide whether they shall explore (choose a meaning for which they do not already know a name) or teach (choose a meaning for which they already know a name). With such strategies, what type of pattern could we expect at the population level? Exploring too much may lead to the introduction of too many conflicting *word-meaning* associations in the population, whereas not enough may lead to really slow completion of the common vocabulary at the population scale, even though the shared core is already stable. However, exploring wisely (following the autotelic principle [21], [22] or following curiosity-driven learning principles based on learning progress [4]) may control the complexity growth of the global vocabulary, and facilitate the convergence towards a stable and complete conventional system. Such a strategy was already suggested in [20].

In what follows, we specify such strategies, present adapted measures to compare them, as well as an experimental comparison in a subset of the parameter space.

II. METHODS

A. Assumptions about Agents and Strategies

As we mentioned earlier, Naming Games exist in a great number of distinct forms. Hence the need to describe which of them we will consider in this work and the assumptions we will make.

An agent's vocabulary will consist in an association matrix – M rows for M available meanings, W columns for W available words – filled with 0s and 1s, initiated with only 0s. In other words, it is implied that the globally available meanings and words are discrete, symbolic and in finite numbers. Homonyms and Synonyms will be automatically removed at each update. This means that a *meaning* can only be associated to a unique *word* (or none), and vice-versa.

At each time step, two agents are randomly selected and engage in one elementary interaction. Agents may have access to a memory of their own past interaction results. Following all those assumptions, designing new strategies is done by setting those two mechanisms:

- The way an agent picks a *word-meaning* association.
- The way an agent remembers the past interactions, by updating its memory.

Those two procedures will be described for each strategy, in the corresponding paragraph II-C.

B. Measures

How do we exactly compare the strategies? How do we quantify the differences between their dynamics? As said earlier, it is usually done using the success rate – i.e. the proportion of successful interactions in a certain time window of the recent past [19], [20]. But the observed success rate itself though is meaningless in the case of a strategy involving the choice of exploration or not. If the common shared vocabulary converges, but to only a subset of the available meanings, the success rate can reach close to 100% if nobody explores, while the vocabulary stays far from complete. In this case the success rate wouldn't indicate anything about convergence towards a complete vocabulary for the entire population. That's why the observed success rate wasn't used in [20], but rather what we can call the *theoretical* success rate, the expected success rate if the strategy would be the naive one (defined at II-C1), in the current state of the population of agents.

Another way of comparing strategies is to actually get the mean time needed for the population to converge to a globally stable shared vocabulary. But this may be really long to compute, especially for non-efficient or even potentially looping strategies – certain sequences of interactions can lead to unchanged vocabularies, if those interactions are repeated indefinitely, the population never reaches convergence. All this doesn't give any indication on a distance of the current state to the fully converged state.

Lastly, the average number of meanings per agent can describe local convergence to a complete vocabulary, but doesn't tell us if the global vocabulary includes a lot of conflicting associations or not, therefore it cannot be used as a global convergence index.

Let's introduce another way of measuring distance to converged state. What is actually such a distance? What does it represent? At the level of the agent, not being in a locally converged state (i.e. a completed vocabulary), means that the agent *doesn't know yet* what the conventions would be in the completed state, among all the remaining

possibilities. In other words, he lacks a certain quantity of information to build a complete vocabulary. What is this missing information, in bits (as defined in [23]), needed by an agent to fill its association matrix? If we have M available meanings and W available words ($W \geq M$), the number of possible configurations for a completed vocabulary is:

$$\Omega = \frac{W!}{(W-M)!} \quad (1)$$

Therefore, the information needed, in bits, to define such a vocabulary among all possible ones is:

$$I(M, W) = \log_2 \Omega = \sum_{k=0}^{M-1} \log_2(W-k) \quad (2)$$

Now, if we consider an intermediate state, in which an agent already has a vocabulary of m associations, without synonyms or homonyms, we can also calculate the information missing in this case. It is quite simple, as the already known meanings and words do not matter anymore in the calculation: By removing them, we get back to the initial problem, i.e. evaluating the missing information of an empty vocabulary, now with $M-m$ available meanings and $W-m$ available words:

$$i(m, M, W) = I(M-m, W-m) = \sum_{k=m}^{M-1} \log_2(W-k) \quad (3)$$

How do we adapt this measure to the entire population? If we average this quantity over the set of agents, it seems equivalent to the mean number of associations, we just applied a logarithm. But we can compute this quantity over the vocabulary actually shared by the population, keeping only associations that spread population-wide. This function is monotonic, and represents a coherent distance to global convergence. Its variations match the order of magnitude of the global dynamics (see for example figure 2), as vocabulary sizes are weighted, through the logarithm, respectively to the probability of transition between them. That correlates the fact that the first population-wide common association will take far more time to appear than the last one, contrary to a *number of links*-type measure, where they will have a really small impact on the representation of the index.

However, this population-wide measure has an issue: we cannot efficiently discriminate a population where all agents but one share the same vocabulary, from a population where all agents have random vocabularies. Clearly the first one is much closer to convergence than the second. How can we build an adapted measure then? By the same approach, but to a different scale. If we average this quantity calculated over all the possible couples of agents, it gives us another quantity, found between the two others, and close to the theoretical success rate. The scale used here is coherent, because interactions happen at this very scale. Other intermediate scales could be used and reveal other properties of the dynamics, but for the purpose of this paper the 2 agents-scale is sufficient. However, it is quite difficult to

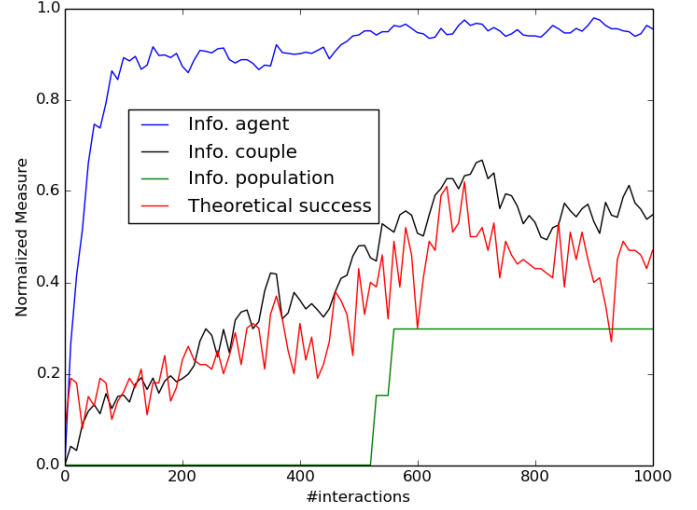


Fig. 2. Comparison of different measures of convergence, normalized, for one simulation with the naive strategy ($M=W=N=10$). Time unit is one speaker/hearer interaction. Blue: Average acquired information per agent, Black: Average acquired information per couple of agents, Green: Acquired information at population level, Red: Theoretical success. The black one describes the best the state of the population (starts at 0, low noise, dynamics almost linear towards globally converged state)

compute, because the number of couples of agents increases as the population size squared. The same problem occurs with theoretical success rate though. Calculations are made by sampling over the set of couples. From this results a non-monotonic behavior and even some noise, which is in fact helpful to guess some other properties of the global dynamics (see II-B).

On figure 2, four different measures are shown, for the same single simulation, following the naive strategy. The acquired information, normalized between 0 (initial state) and 1 (maximum, population in fully converged state) at 3 different scales: averaged for all agents (in blue), sampled over couples of agents (black), and taken on the whole population (green). The fourth measure is the theoretical success rate (here equal to observed success rate, in red). The green measure is monotonic, but it represents badly the overall dynamics, because too less associations are taken into account. The blue measure grows too quickly compared to the overall dynamics, and takes too many associations into account. The red measure doesn't start at 0 (random initiation interactions may lead to success, with a probability of $\frac{1}{W}$, 10% in this case), is quite noisy and heavy to compute, but seems still a good indicator as its dynamics matches the overall variations. The black measure is close to the red one, hence has the its advantages, but is less noisy, starts at 0, and is meaningful from an information theoretical point of view. Also, it cannot go under the green one, whatever the cardinality of the sample is, as it represents the theoretical minimum, for all possible couples of agents. This is not the case for the theoretical success, as can be seen in figure 2.

In the remaining part of the paper, the normalized common acquired information of couples of agents will be used as a measure to compare strategies.

It can be noted that (low) noise and non-monotony are useful features ; in fact it indicates that the population is still out of equilibrium, what allows us when the dynamics is slow to discriminate a blocked state – with a stable but not complete vocabulary – from a high-competition state where overall dynamics are expectedly slow.

C. Strategies

Now that we can quantify features of the strategies, let's introduce some of them.

We'll start by introducing the naive strategy, which is the standard strategy usually used in Naming Games, and also give a proof of its convergence. Next, we will introduce active learning strategies: first the *Success-Threshold* type of strategies, as described in [20]. Since those strategies depend on a parameter, we will suggest a method to find an optimum for this threshold-parameter. Another strategy, which we will call *Last Result* strategy, can be seen as a variant of the Success-Threshold. Eventually, we will define a general type of strategies with an active learning behavior, and suggest interesting ones based on calculations. Comparison of those strategies can be found in section III-A.

Algorithm 1 NAIVE STRATEGY

The two mechanisms needed to describe a strategy are detailed below: how the enacted association is chosen during an interaction, as well as the update of both speaker and hearer after the interaction, involving the chosen association, as well as the meaning guessed by the hearer. Both agents add the enacted association to their vocabulary, and remove any conflicting synonym/homonym.

procedure PICK LINK (vocabulary)

$m \leftarrow$ get random meaning

$w \leftarrow$ get word(m)

return (m, w)

end procedure

procedure UPDATE (m_s, w, m_h)

add link (m_s, w)

remove homonyms (m_s, w)

remove synonyms (m_s, w)

end procedure

1) *The Naive Strategy*: What we call the naive strategy is the strategy usually used in Naming Games instances as a standard [16], [19], [20]. The meaning-choice policy is simply uniform, i.e. meanings are equiprobable, whatever their state, associated to a word or not. As far as post-interaction vocabulary update is concerned, usually synonyms and homonyms are used and help a lot in the convergence process (see 9). In fact, they carry information about the past interactions and help to smooth the competition between associations. In our models we will only consider the no-homonyms no-synonyms version of the naive strategy (algorithm 1), however in III-B the actual impact of this assumption will be discussed.

An interesting remark: this strategy leads to convergence. It can be proven easily that the probability to be in a converged state at infinite time is equal to one. There exists a combination of $N - 1$ interactions in which an agent teaches one special association to all other agents in the population (of size N), among a finite (but extremely big) number of possible sequence of $N - 1$ interactions. Therefore the probability to get a fixed association population wide after a certain number of steps is finite and non-zero. If M is the number of available meanings, this reasoning repeated M times leads to convergence at a non-zero finite probability after a long enough time. At infinite time, this probability eventually becomes one. As a conclusion, the naive strategy almost surely reaches convergence.

Algorithm 2 SUCCESS-THRESHOLD STRATEGY

Explore() means choosing an unknown meaning, and teach() choosing a known one. Success rate is computed as the proportion of average ratio $S/(S+F)$ over the known meanings in the speaker's vocabulary. S and F respectively stand for success and fail. *threshold* is by default set at 90%

procedure PICK LINK (vocabulary, S, F)

if success rate(*known meanings*) > *threshold* **then**
 explore()

else

 teach()

end if

return (m, w)

end procedure

procedure UPDATE (m_s, w, m_h)

add link (m_s, w)

remove homonyms (m_s, w)

remove synonyms (m_s, w)

if $m_s = m_h$ **then**

$S[m_s] \leftarrow S[m_s] + 1$

else

$F[m_s] \leftarrow F[m_s] + 1$

end if

end procedure

2) *Success-Threshold Strategy*: The type of strategies developed in [20] is a first approach to active learning policies. The main idea is to choose exploring new meanings over teaching known ones if the success rate over the interactions involving the already associated meanings exceeds a certain threshold, in the original work set at 90% (see algorithm 2). From a cognitive point of view, the agent seeks to learn new things only when confident enough about what it previously learned. This policy actually led to interesting results, especially convergence time proportional to population size, when exponential for the naive strategy.

What we wonder now is, is this threshold of 90% wisely chosen, or would another value give better results? To answer this, we will compare strategies following the same policy, but with a different value of the threshold. Convergence time may reach high values, or in certain cases not even

exist, the shared information measure over sampled couples of agents will here be really useful. But because we want to compare a large number of different strategies (threshold varying continuously in $[0, 1]$), representing the evolution of this index for all of them is not convenient. Instead, discretizing time by representing snapshots of index values as a function of this threshold is way more convenient, and the most efficient threshold values appear clearly.

The results, for $M=W=N=20$, are represented on figure 3. When the low values (almost only exploration) stay far from convergence, values above 80% for the threshold parameter lead to fast convergence. For 100%, agents never explore because they wait for their success rate to reach this value, situation which in our case (success rate computed over all past interactions) cannot be reached after a single failed interaction. Agents never explore, hence never get a complete vocabulary. Values close to 100% exhibit a similar pattern. According to all this data, 90% seems close to an optimal choice, at least for the chosen values of M , W and N .

Algorithm 3 LAST RESULT STRATEGY

A variation of the Success-Threshold strategy

```

procedure PICK LINK (vocabulary, success)
  if success = 1 then
    explore()
  else
    teach()
  end if
  return ( $m, w$ )
end procedure

```

```

procedure UPDATE ( $m_s, w, m_h$ )
  add link ( $m_s, w$ )
  remove homonyms ( $m_s, w$ )
  remove synonyms ( $m_s, w$ )
  if  $m_s = m_h$  then
    success  $\leftarrow$  1
  else
    success  $\leftarrow$  0
  end if
end procedure

```

3) *Last Result Strategy*: What we call here the *Last Result* strategy is a simplification of the previous one, the time window of the recent past taken into account being reduced to its minimum (1 interaction). In other words, we consider only the last interaction and its result, success or fail. Success leads the agent to explore, whereas previous failure makes it teach a known meaning. The value of the threshold parameter doesn't matter here, because taken on one interaction the success rate it is compared to can only be either 0% or 100%. See algorithm 3. From a cognitive point of view, this strategy is a limitation to extremely short-term memory.

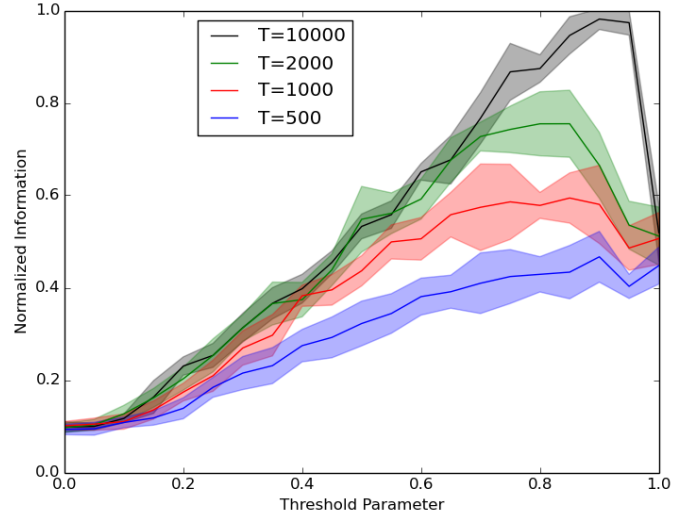


Fig. 3. Convergence speed comparison according to threshold parameter, for the success-threshold strategy. The previously selected value of 90% is well-chosen. $M=W=N=20$, 8 iterations. Time unit is number of past interactions.

Algorithm 4 DECISION VECTOR STRATEGY

Exploration is triggered if a 1 is present in the decision vector at the coordinate corresponding to lexicon size

```

procedure PICK LINK (vocabulary)
  if decisionvector[#(vocabulary)] = 1 then
    explore()
  else
    teach()
  end if
end procedure

```

4) *Decision Vector Strategies*: By getting rid of any memory of the past interactions other than the lexicon itself, and remembering that the absence of synonyms/homonyms reduces the configurations to equivalent permutations at fixed number of associations, one degree of freedom remains for the definition of a strategy: the choice of whether to explore an unknown meaning or teach a known one. In our models this is a function of only the number of meanings already known, in other words the behavior can be represented by a vector of dimension $M + 1$, where M is the total number of available meanings, and each coordinate would be the probability to explore, if found as a speaker with this given number of known meanings. This is detailed in algorithm 4.

The simplest form of this type of strategy would be described by a vector of elements only equal to 0 or 1. That means, choose to explore only at certain given vocabulary sizes, and teach when at others. From a cognitive point of view, it can be seen as a maturational process. Certain states trigger exploration, which is inhibited otherwise.

An important remark at this point is that it may result in a stable non-converging state. For example, if we choose to put only zeros on this vector (excepted for the first value, that has to be 1), the common shared vocabulary will not

include any other associations than the ones corresponding to initialization of agents' vocabularies. This would be at maximum $N-1$, and on the average around $N/2$. Then those associations may be conflicting, but still the maximum size of the final vocabulary is $N-1$, and in the case $M=W=N$, it means that the common shared vocabulary cannot converge to a complete one (size N).

A first approach would be to evaluate the approximate size of this vocabulary, and set to 1 the corresponding coordinate of the decision vector, keeping 0s before. Exploration would be triggered at this very vocabulary size, and not for an inferior value. As mentioned in the previous sentences, we expect this value to be a bit less than $N/2$. Let's assume that the ratio keeps being the same between two steps of this evolution. In other words, if we consider this ratio to be exactly 2, we would put ones in the decision vector at the coordinates $N/2$, $3N/4$, $7N/8$, etc. If the ratio is r , 1s would be found at coordinates $N(r-1)/r$, $N(r^2-1)/r^2$, $N(r^3-1)/r^3$, etc. This *geometrical* pattern is graphically described in figure 5.

In figure 4, we tested this hypothesis for different values of r , and the behavior was quite interesting: in all cases we do not reach full convergence quickly (compared to other studied strategies, see figure 8) but we can distinguish those which will fully converge from those that won't, by observing the local behavior (noisy or not), following the reasoning detailed in paragraph II-B. If there is no variation at all, it means that all agents share the exact same vocabulary. As it is the type of behavior we expect with an all-0 decision vector, it means that the first 1 in the vector is higher than the expected equilibrium vocabulary size in this all-0 case, and the ratio r chosen is too high. On the contrary, variations would indicate non-equilibrium and high meaning competition, and probable convergence after. In other words, agents would have explored new meanings too early, the ratio r is too low.

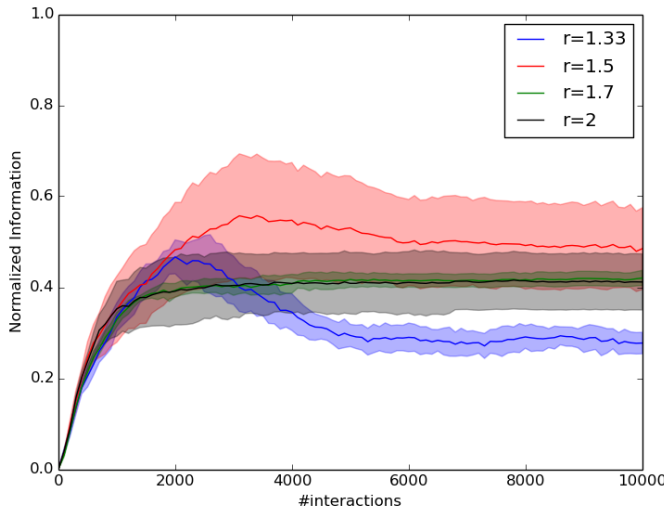


Fig. 4. Geometrical decision vector strategies. Two behaviors are observed: stable sub-converged state (ratio > 1.6) or high meaning competition (ratio < 1.6). $M=W=N=20$, 4 iterations.

As we can observe on figure 4, at a certain value (1.6) of r a pattern change can be observed. It would be interesting to do a deeper study of this phase transition, especially the dependence on M , W and N parameters, but it goes beyond the scope of this paper.

$$\begin{array}{c} \overbrace{\overbrace{1 \ 0 \ 0 \ 0}^{L_1} \ 1 \ 0 \ 1 \ 1 \ 0}^{L_2} \\ \left[\begin{array}{cccccccc} 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{array} \right] \quad \frac{L_i}{L_{i+1}} = r \\ m = \quad 0 \ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \end{array}$$

Fig. 5. Pattern for the geometrical decision vector, with $M = 8$ and $r = 2$. An agent explores when possessing 0, 4, 6 or 7 associations.

5) *Gain Maximization Decision Vector*: Another approach, staying in this scope of strategies, is to explore until a certain vocabulary size, and then wait for the introduced competing associations to mix into a common vocabulary. The change of behavior shall occur at a wisely chosen vocabulary size, otherwise it may converge really slowly or not converge at all.

But how to choose this value? A first guess could come from an approximation of expected information gain by the speaker. Here, we will only consider expected information gain as a hearer at the next step. Information gained through exploration as a speaker would heavily bias the policy towards exploration, because it corresponds to a selfish behavior not leading at all to fast convergence. This could be understood as optimizing the acquired information by the speaker only, and not both agents.

So what would be this expected gain of information? As a hearer, an agent can either gain a link (the one enacted by the speaker is brand new and not in conflict with the existing ones), lose one (enacted association conflicting word-wise and meaning-wise), or stay at the same lexicon size (only one type of conflict). As the change in behavior will appear at the end of an all-exploration phase, we will consider that associations encountered in other agents are initiated completely separately in this first phase, i.e. that lexicons are independent of each other. This assumption is actually coherent considering the shape of the resulting vector, as the all-exploration phase takes place only at the beginning.

Knowing all this, we can compute the expected gain referring to the table in figure 6. For each possible outcome, probabilities are calculated for a totally random association (all MW possibilities equiprobable). It means speaker and hearer share at this point no information at all, which is in our case coherent, as exploration is done from the beginning. For example, the probability of getting +1 association for the hearer possessing m associations in its vocabulary is calculated this way: $(M-m)$ meanings are unused, as well as $(W-m)$ words. Therefore $(M-m)(W-m)$ associations among all possible give the expected outcome (+1). The probability, as all associations are equiprobable, is the ratio between those two numbers. The resulting vector shape is represented in figure 7.

| Associations | Info. Gain $G_m = \Delta i$ | Probability p_m |
|--------------|-----------------------------|--------------------------|
| +1 | $\log_2(W - m)$ | $\frac{(M-m)(W-m)}{MW}$ |
| 0 | 0 | $\frac{m(1+M+W-2m)}{MW}$ |
| -1 | $-\log_2(W - m + 1)$ | $\frac{m^2 - m}{MW}$ |

Fig. 6. Hearer’s possible outcomes with associated gains and probabilities, when having m associations in the vocabulary. M and W stand for total available meanings and words. Probabilities are calculated for a random association (uniformly over all meanings and words) chosen by the speaker.

$$\begin{array}{c}
\mathbf{E}(\Delta i) > 0 \quad \Big| \quad \mathbf{E}(\Delta i) < 0 \\
\left[\begin{array}{cccc|cccc} 1 & 1 & \cdots & 1 & 0 & \cdots & 0 & 0 \end{array} \right] \\
\mathbf{E}(\Delta i) = \sum_{\alpha \in \{-1,0,1\}} p_m(\alpha) \cdot G_m(\alpha)
\end{array}$$

Fig. 7. Pattern for the Gain Maximization decision vector. An agent explores when expected information gain (as defined by the given formula, using the table from figure 6) as a hearer is positive. For the chosen range of parameters, the separation is roughly located at the middle of the vector.

III. RESULTS AND DISCUSSION

A. Comparing Strategies

Figure 8 shows the evolution of the acquired information through time for different strategies. For each strategy, the quantity is averaged on several iterations, and standard error is represented. In all cases, number of available meanings, words, and size of the population are all equal to 20. We represented the naive strategy, as well as several active learning strategies: the *Last Result* strategy II-C3, the *Success-Threshold* strategy II-C2, and the *Gain Maximization* decision vector strategy II-C5.

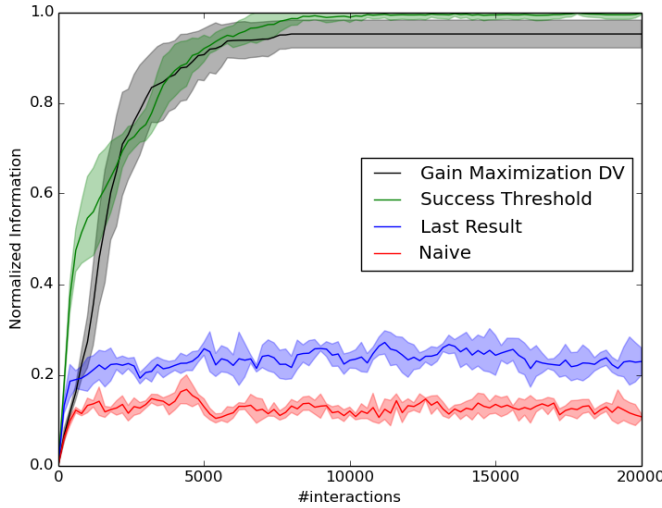


Fig. 8. Comparison of different strategies. Active learning ones converge faster than the naive one. $M=W=N=20$, 4 iterations

The naive strategy stays at low values of acquired information, indicated too much exploration and high meaning-competition among the population. It will eventually fully converge, but its dynamics are really slow.

The Last Result strategy performs a bit better than the naive strategy, but also presents quite slow dynamics compared to the other two curves.

The Success-Threshold strategy converges really fast, as well as the decision vector strategy. It can be noted that the decision vector strategy doesn’t fully converge, and stays blocked, but still reaches as fast as the Success-Threshold strategy really high values of acquired information. The standard error also indicates that in some cases it actually converges, and would be in those particular cases (not in general) the fastest converging strategy. Otherwise the Success-Threshold strategy is.

These observations support a significant acceleration of the global dynamics carried by the regulation of complexity growth, even if only a subset of those strategies have been studied, and more efficient ones may exist. Also, it has to be noted that those conclusions are drawn from the study of only a subset of the parameter space, and deeper exploration would be needed. But still, it shows that active learning can have a significant impact.

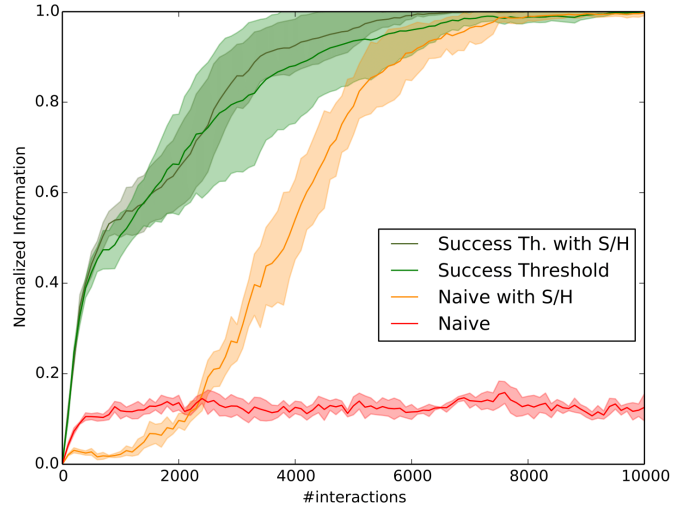


Fig. 9. Influence of use of synonyms and homonyms. It increases a lot the naive strategy’s convergence speed, but doesn’t change significantly the success-threshold strategy’s pattern, which converges faster in both cases. $M=W=N=20$, 4 iterations.

B. Influence of the Main Assumption over Synonyms and Homonyms

The main difference between our assumptions and the usual Naming Games is the absence of homonyms and synonyms. Even if the measure isn’t as well-defined for such vocabularies as under our assumption, we can adapt it and get a clue on the state of such a population, and compare the dynamics to our models. The reason why it isn’t considered as being well-defined is because we cannot consider shared associations the same way. If a given association is shared

at the population level, it may as well disappear later, as having synonyms or homonyms which would overcome this very association in the final globally converged vocabulary. The adaptation used here is calculation of the acquired information at the level of couples of agents, but using only shared associations which don't have any synonym or homonym. Clearly, this approximation of the acquired information is underestimated, as synonyms and homonyms actually carry some information. However, at low level of synonymy and homonymy it fits the approximated quantity, which is the case in the last part of the dynamics (see [16])

The question now is, if homonyms and synonyms carry some information, would active learning strategies benefit from their use or not at all? Figure 9 compares the naive strategy with the Success-Threshold strategy, both with and without synonyms/homonyms.

The naive strategy, as expected, benefits greatly from the use of synonyms and homonyms, and converges in the represented time window only in this case. Without synonyms and homonyms, convergence would be reached several order of magnitudes further, as can be guessed by the shape of the curve. For the active learning strategy however the curves almost mix, which implies that synonyms and homonyms are almost unused. To rephrase it, the information conveyed by synonyms and homonyms in the naive strategy may be already partly gained by the active learning choice policy. Just a small difference in acquisition of the last associations is observed, but this strategy is still in all cases faster converging than the naive one.

C. Future Work

We wanted to study the impact of developmental/active learning approaches on Naming Games' convergence speed. In order to do that, we introduced a new measure of distance to the converged state, the normalized acquired information shared by couples of agents, based on information theory principles. The newly introduced strategies show some interesting patterns in the Naming Games dynamics, in particular they converge faster than the naive strategy. Also, the information carried by the active control of complexity can be enough to avoid using synonyms and homonyms.

This work could be extended to a systematic study of strategies, especially dependence in M , W and N (the parameter space not having been explored significantly for the moment), after a complete inventory of possible strategies. Then, characterizing the flow of information could help to conduct a deeper study of the Gain Maximization approach, and maybe derive a theoretically optimal strategy. Eventually, other Language Games, like the Guessing Game [12] or the Categorization Game [16], could be analyzed in the same way.

ACKNOWLEDGMENT

We thank everyone who provided feedback, especially Jonathan Grizou. This research was partially funded by ERC Grant EXPLORERS 240007, and also École Normale Supérieure de Lyon in the form of a 4th year study project.

SOURCE CODE

The code used for the simulations of this paper was written in Python. It is available as open source software, along with explanatory notebooks, on the Inria Flowers team github: <https://github.com/flowersteam/naminggamesal>

REFERENCES

- [1] L. Steels, F. Kaplan, A. McIntyre, and J. Van Looveren, "Crucial factors in the origins of word-meaning," *The transition to language*, vol. 12, pp. 252–271, 2002.
- [2] P.-Y. Oudeyer, *Aux sources de la parole: auto-organisation et évolution*. Odile Jacob, 2013.
- [3] A. Cangelosi, G. Metta, G. Sagerer, S. Nolfi, C. Nehaniv, K. Fischer, J. Tani, T. Belpaeme, G. Sandini, F. Nori *et al.*, "Integration of action and language knowledge: A roadmap for developmental robotics," *Autonomous Mental Development, IEEE Transactions on*, vol. 2, no. 3, pp. 167–195, 2010.
- [4] P.-Y. Oudeyer, F. Kaplan, and V. V. Hafner, "Intrinsic motivation systems for autonomous mental development," *Evolutionary Computation, IEEE Transactions on*, vol. 11, no. 2, pp. 265–286, 2007.
- [5] C. Moulin-Frier and P.-Y. Oudeyer, "Exploration strategies in developmental robotics: a unified probabilistic framework," in *Development and Learning and Epigenetic Robotics (ICDL), 2013 IEEE Third Joint International Conference on*. IEEE, 2013, pp. 1–6.
- [6] C. Moulin-Frier, S. M. Nguyen, and P.-Y. Oudeyer, "Self-organization of early vocal development in infants and machines: the role of intrinsic motivation," *Frontiers in psychology*, vol. 4, 2013.
- [7] P.-Y. Oudeyer, "The self-organization of speech sounds," *Journal of Theoretical Biology*, vol. 233, no. 3, pp. 435–449, 2005.
- [8] P.-Y. Oudeyer and L. Smith, "How evolution may work through curiosity-driven developmental process," *Topics Cogn. Sci.*, 2014.
- [9] J. Gottlieb, P.-Y. Oudeyer, M. Lopes, and A. Baranes, "Information-seeking, curiosity, and attention: computational and neural mechanisms," *Trends in cognitive sciences*, vol. 17, no. 11, pp. 585–593, 2013.
- [10] F. Kaplan, P.-Y. Oudeyer, and B. Bergen, "Computational models in the debate over language learnability," *Infant and Child Development*, vol. 17, no. 1, pp. 55–80, 2008.
- [11] M. Lopes and L. Montesano, "Active learning for autonomous intelligent agents: Exploration, curiosity, and interaction," *arXiv preprint arXiv:1403.1497*, 2014.
- [12] L. Steels, "Language games for autonomous robots," *Intelligent Systems, IEEE*, vol. 16, no. 5, pp. 16–22, 2001.
- [13] —, "Evolving grounded communication for robots," *Trends in cognitive sciences*, vol. 7, no. 7, pp. 308–312, 2003.
- [14] V. Loreto and L. Steels, "Social dynamics: Emergence of language," *Nature Physics*, vol. 3, no. 11, pp. 758–760, 2007.
- [15] L. Steels, *The Talking Heads Experiment, In Words and Meanings, Vol. 1.*. Special pre-edition for LABORATORIUM, Antwerpen, 1999.
- [16] V. Loreto, A. Baronchelli, A. Mukherjee, A. Puglisi, and F. Tria, "Statistical physics of language dynamics," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2011, no. 04, p. P04006, 2011.
- [17] F. Tria, V. Loreto, V. D. P. Servedio, and S. H. Strogatz, "The dynamics of correlated novelties," *Scientific reports*, vol. 4, 2014.
- [18] R. Ferrer i Cancho and A. Díaz-Guilera, "The global minima of the communicative energy of natural communication systems," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2007, no. 06, p. P06009, 2007.
- [19] F. Kaplan, "La naissance d'une langue chez les robots," 2001.
- [20] P.-Y. Oudeyer and F. Delaunay, "Developmental exploration in the cultural evolution of lexical conventions," in *proceedings of the 8th international conference on epigenetic robotics: modeling cognitive development in robotic systems*. lund university cognitive studies, 2008.
- [21] L. Steels, "The autotelic principle," in *Embodied Artificial Intelligence*. Springer, 2004, pp. 231–242.
- [22] L. Steels and P. Wellens, "Scaffolding language emergence using the autotelic principle," in *Artificial Life, 2007. ALIFE'07. IEEE Symposium on*. IEEE, 2007, pp. 325–332.
- [23] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 1, pp. 3–55, 2001.